DOCUMENT RESUME

ED 261 064 TM 850 455

AUTHOR Martin, David W.
TITLE Applied vs Basic Research: On Maintaining Your Balance with a Foot in Each Camp.
PUB DATE Apr 85
NOTE 17p.; Presidential Address presented at the Annual Meeting of the Rocky Mountain Psychological Association (Tucson, AZ, April 24-27).
PUB TYPE Speeches/Conference Papers (150) -- Viewpoints (120) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Action Research; Computer Simulation; Educational Experiments; *Educational Research; *Flight Training; Higher Education; Learning Strategies; Psychological Studies; *Research Design; Research Methodology; *Research Problems; Research Utilization; *Training Methods; Transfer of Training; Visual Stimuli

ABSTRACT

The paper discusses a number of issues concerning the practical usefulness of cognitive psychology research, and presents a case study of pilot training methods to illustrate a model of research processes that produces outcomes which contribute to both basic and applied research goals. Research studies are described as varying in the degree to which they include reviews of the literature, present practical conclusions, select and manipulate variables based on convenience or bias, or add to the existing body of knowledge. The following research strategy is recommended: identify application set; determine common dimensions; create complex generic laboratory task to test the theoretical hypothesis; perform experiment; form theoretical conclusion; validate conclusion and task; and list applications. This strategy is illustrated in detail with a case study. The case study describes flight training of aircraft pilots, with emphasis on their attention to multiple visual stimuli. (GDC)

ED261064

Applied vs Basic Research: On Maintaining Your Balance With a Foot in Each Camp

David W. Martin

TM 850 455

Applied vs Basic Research: On Maintaining Your Balance With a Foot in Each Camp

David W. Martin

Before beginning my talk I would very much like to thank the members of Rocky Mountain Psychological Association for the privilege of serving as your President this past year. Our Association is very fortunate to have an experienced and dedicated executive committee that makes the President's job relatively easy. I do believe that the Association is healthy and that we continue to retain our unique character as an informal colleagial association. I am sure that Nelson Jones will do a fine job this next year in leading the Association.

Many of the past presidents of Rocky Mountain Psychological Association have used the Presidential Address as an opportunity to summarize the research trends in their careers. I considered whether this would be appropriate in my case. I decided that attempting to talk about decision making, organization in memory, dual task performance, and models of attention in one talk would have tried the patience of even my best friends. For this reason I have decided to talk about research strategies. In particular I am concerned about the integration of applied and basic research into a common theoretically structured body of knowledge and I feel that the research models currently used in basic and applied research minimize the possibility of such integration. What I will do today is show you a model that I have been using for research. I believe that this model provides a means for structuring research processes so that the outcome contributes to both basic and applied goals. I will illustrate the model by telling you about a case study currently underway in my laboratory at New Mexico State University.

In the suggested model you will undoubtedly see my engineering psychology bias. However, I believe that the model could be applied to other areas of research in which an attempt is made to find answers that are of use in application areas, particularly areas of research that use simplified laboratory paradigms.

As a graduate student at Ohio State University I was taught and convinced of the effectiveness of a research strategy whereby good basic theoretical research should provide answers that help solve applied problems. Yet, I see little of that happening, especially in the area of basic cognitive psychology. Instead I see numerous papers using Stroop color naming, Posner letter matching, Sperling partial report, Sternberg additive factors, Schvaneveldt priming, and lexical decision tasks as well as the traditional paired-associates and serial learning of nonsense syllables.

Attempting to derive principles that are useful for solving realworld problems from research based on such simplistic paradigms is exceedingly difficult. To illustrate this problem, back in the early seventies I was hired as a consultant to extract principles from the basic research literature. These principles were meant to help systems designers construct easy to use codes and abbreviations for use in computerized information systems. At first this task appeared to be straightforward. Think of the sheer volume of research done using nonsense syllables. Nonsense syllables would seem to be quite similar to codes or abbreviations. The problem is that

APPLIED RESEARCH

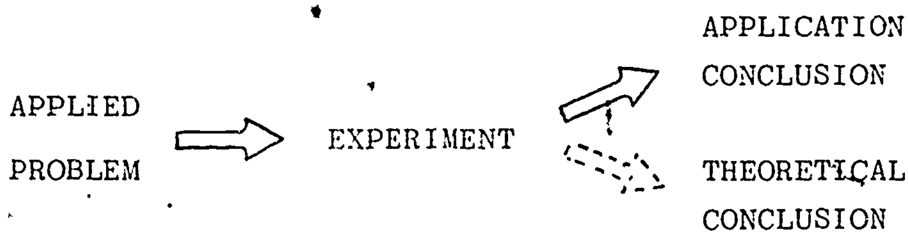


Figure 1. Research strategy employed by most researchers working in applied settings.

BASIC THEORETICAL RESEARCH

THEORY ⇨ PARADIGM ⇨ EXPERIMENT ⇨ THEORETICAL CONCLUSION ⇢ APPLICATION

Figure 2. Research strategy employed by basic theoretical researchers.

nonsense syllables, by definition, have no meaning. What we were looking for were principles by which codes could be constructed which conveyed the most meaning. In the end, very little of the basic research literature was of much use for this purpose.

Others who have attempted to use the basic literature to derive practical principles have been equally frustrated. Chris Wickens in his new book Engineering Psychology and Human Performance makes a valiant attempt to do this. Yet in a letter to the editor of the Human Factors Society Bulletin, David Meister criticized the book as failing to be a "textbook on human factors." Meister goes further in saying that, "It is increasingly apparent to some of us that there are two segments of our discipline: one, that of the behavioral researcher, the other, that of the human factors practitioner. These worlds rarely intersect, and the researcher's interest in application... is almost nonexistent."

The exception to this pattern seems to be motivated more by a change in research funding than by a perceived need on the part of academic researchers to work in the barren wasteland between basic and applied research. While I have heard many of my colleagues lamenting the cutback in federal funds for basic research, I believe that the swing in funding at the federal level to the Department of Defense and the increased involvement of corporations in funding academic research has had one positive effect: forcing some very good basic researchers to consider research strategies that make applied as well as basic contributions. Some of these people are now using research strategies similar to the one that I will suggest today. However, the large majority of researchers still seem to be using either the basic or applied strategies that I will present next.

Figure 1 shows the relatively simple strategy employed by most researchers working in applied settings. Essentially an applied problem presents itself, the researcher designs an experiment within the context of this applied problem, carries it out, and at the conclusion of the experiment makes a recommendation aimed at solving the original applied problem. In a few cases, the researcher has studied the literature sufficiently to bring theory to bare on the problem and can make a theoretical conclusion as well as an application conclusion. However, with the exception of a few applied settings where researchers are encouraged to publish experimental results, the payoff for taking the extra effort required to reach a theoretical conclusion is so small that researchers seldom do so. Usually the applied researcher is concerned with comparing two human-machine systems or two instructional modes or two therapeutic techniques in order to find out which one is superior. The goal is simply to make a choice, not to develop a theoretical structure that might add to the basic body of knowledge.

Figure 2 shows a model which I believe describes much of the basic theoretical research done today. In this case, the researcher starts with a theoretical question generated from the body of knowledge. The researcher then develops a simple experimental paradigm which allows theories to be pitted against one another. Next the researcher designs an experiment that manipulates the variables of interest and draws the appropriate theoretical conclusion from the results of the experiment. Most basic researchers stop here because they have answered the original theoretical question and can update the body of knowledge. A few researchers will generalize their

MOST ACADEMIC RESEARCH

PARADIGM ⇨ THEORETICAL DICHOTOMY ⇨ EXPERIMENT ⇨ THEORETICAL CONCLUSION

Figure 3. Research strategy employed by most academic researchers.

SUGGESTED RESEARCH STRATEGY

IDENTIFY APPLICATION SET ⇒ DETERMINE COMMON DIMENSIONS ⇒ CREATE COMPLEX GENERIC LAB TASK ⇒ ... ⇒ VALIDATE CONCLUSION & TASK

APPLICATION 1, APPLICATION 2, ..., APPLICATION n ⇒ DIMENSION 1, DIMENSION 2, ..., DIMENSION n ⇒ TASK ⇒ THEORETICAL HYPOTHESIS ⇒ EXPERIMENT ⇒ THEORETICAL CONCLUSION ⇒ APPLICATION 1, APPLICATION 2, ..., APPLICATION n

Figure 4. Research strategy that combines the applied and basic models.

conclusion to one or more applications but this step is rarely done. Indeed, if a simple laboratory paradigm has been used, generalizing to the more complex application setting is often tenuous at best.

I would argue that the model shown in Figure 2 characterizes the sequence of events for the best academic researchers, but in reality most academic research follows the model shown in Figure 3. In this model, some leader in the field, or perhaps the student's advisor, has developed a simple laboratory paradigm which the researcher takes as a starting point. The researcher then decides to manipulate one of the parameters available in the paradigm, often more for reasons of convenience (e.g. the equipment is already available, the software has been developed, the methods have been validated) rather than for theoretically justifiable reasons. The researcher then looks for a post hoc theoretical reason which justifies the importance of this manipulation and usually pits two theories against one another setting up a theoretical dichotomy. The experiment is conducted and the results support one of these theories while refuting the other. These reseachers are engaging in what they consider to be an interesting intellectual game which seldom provides any useful application information. Indeed, I would argue that such research seldom generates information that is of much use for building the body of knowledge. Judging from the literature, the theoretical dichotomy often turns out to be a false dichotomy better represented by two points on a continuum. Once the continuum has been recognized, the research becomes considerably more complex and it loses its fascination as an intellectual game. These researchers then move on to a new game.

I would like to suggest a research strategy that permits us to combine the best of the applied and basic models I have just presented. I hope that the strategy I am going to suggest will allow researchers to conduct studies that provide answers to applied problems yet which also contribute to the longer lasting theoretically structured body of knowledge. In my laboratory we have found this strategy to be useful.

Figure 4 shows the suggested research strategy. As you can see, the first step is to identify an application set. That is, the researcher attempts to find a number of real-world applications that seem to have structural commonalities. From this set of applications the researcher next determines what common dimensions characterize this application set. Once the dimensions are determined, the researcher attempts to create a generic laboratory task that contains most of these dimensions. Within the constraints imposed by the set of dimensions, the researcher attempts to make the laboratory task as simple as possible, but the number of dimensions will generally require the generic laboratory task to be relatively complex compared to the simple paradigms currently in use. These first three steps are the critical part of this research strategy. Perhaps they are as much art as they are science. They certainly require good guesses on the part of the researcher. Yet, as psychologists, these steps are not foreign to us. They require the skills that psychologists possess: observation, analysis, and synthesis.

Once the generic laboratory task has been created, it can be used as a testbed for many theoretical hypotheses. The initial experimentation using the generic task is more difficult than using a simple paradigm. Considerably more pilot experimentation is necessary to "tune" the task. Ceiling and floor

levels must be determined and the complexity of the task must simulate the complexity of the application set. Even after the generic task meets these requirements, experimentation is still more time consuming than using a simple paradigm. The one hour of experimentation which is sufficient for a Stroop color naming task will not be sufficient for the type of generic laboratory task suggested here.

The result of experimentation will be a theoretical conclusion, but hopefully one that will also have external validity since the task was designed to simulate the application areas to which it will be generalized. We do not have to guess whether generalization is appropriate, however, since the final step is to validate the theoretical conclusion in one or more application settings. Successful validation will do three things. It will strengthen our confidence in the theoretical conclusion itself. Second, it will increase our confidence in the generalizability of the conclusion. Third, it will strengthen our confidence in the usefulness of the generic laboratory task as an appropriate testbed for the application set.

Now I would like to illustrate the suggested research strategy using one case study from my laboratory at New Mexico State University. We have been conducting this research program for approximately the past five years. A problem that I have been interested in for some time is how people choose to allocate their attention to the multiple sources of information they are exposed to at any given time. In particular, my interest was in the operator of a human-machine system and how that operator chooses which source of information to pay attention to and which sources to ignore. Increasingly such systems are becoming automated and the operator's major function is to monitor system status. While a number of theories of attention have been proposed to explain how humans operate in a multi-task environment, system constraints often force the operator to a single-channel mode. That is, since the sources of information are usually presented visually at various spatial locations and since the operator can look in only one direction at a time, the operator can be considered similar to a single-channel processor. The various sources of information are sequentially processed through this channel.

Given this general characterization of the application area of interest, which applications would seem to fit into an appropriate application set? One application is the nuclear power plant control room. In general, the operator's task is to monitor the displays to determine system status and to occasionally adjust a control to maintain the status. Another application is the aircraft pilot. The flight instrument panel of an aircraft has a large number of spatially separated instruments some of which are considerably more important than others. I flew in the cockpit of a small executive aircraft the other day, and with the exception of the takeoff and landing, the pilot's main task was that of monitoring instruments and occasionally re-programming the microprocessor which flys the aircraft. Other applications that would seem to fit into the application set include air traffic control and many modern industrial processes such as petroleum refineries.

Now that we have identified an application set, what seem to be the dimensions common to this application set? The most obvious characteristic is that there are a large number of spatially separated sources of information. Some of these sources change state at a rapid rate and others at a slow rate.

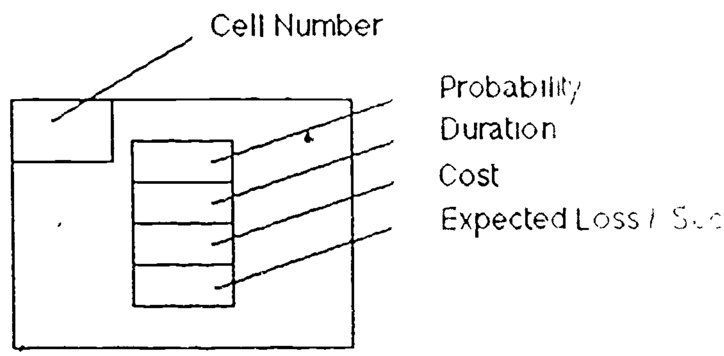
Some sources have a relatively high frequency of targets which require responses and others have a low frequency of targets. Some targets require responses that are relatively difficult and others responses that are easy. Finally, if the operator fails to respond to a target, the cost to the system can be relatively small or disastrous.

Consider the basic six instruments in an aircraft to illustrate some of these dimensions. Various instruments can become more or less important depending upon the particular phase of flight. For example, a pilot flying on instruments who make an instrument landing system approach would have to constantly monitor the attitude display indicator. However, a pilot making a rather steep climb would be more likely to monitor the airspeed indicator to insure that the aircraft was well above stall speed. Some instruments change rather quickly (e.g., the bank indicator during a cross wind). Other instruments change rather slowly (e.g., amount of fuel in the tanks). Responses to targets occurring on some instruments are easy while others are more difficult. Failure to respond to some targets is not very costly (e.g., fine tuning of the fuel mixture) while for other targets it is quite costly (e.g., running out of fuel or reaching stall speed).

From these dimensions we created a generic laboratory task. On the video display of a TERAK computer terminal six single digit numbers are displayed in a 2 by 3 matrix. The six cells in this matrix are similar to the six instruments viewed by the pilot. The single digit numbers are replaced by other numbers according to a predetermined rate. The subject's task is to respond to target numbers. For purposes of this paper, lets consider the targets to be the numbers 1 through 4. When the subject sees a target number, a response is made by pressing one of the six cell location keys with the right hand and then pressing the appropriate target number key with the left hand. No response is required to nontarget digits. Failure to respond to a target digit before it is replaced results in a loss of points. At the beginning of each trial a thousand points is displayed just under the matrix. When points are lost they are immediately subtracted from the running total on the display.

A task is defined entirely by the specification of a parameter set. In Figure 5 you see a parameter set. Three parameters define each cell. By way of example look at cell number 3. The probability that any digit displayed will be a target is .4. You can see that targets are four times more likely in cell 3 than in cell 6. The duration of a digit and correspondingly the rate by which digits are displayed is 2 sec. Thus, each digit is displayed for only 2 sec. in cell 3 and for 4 sec. in cell 5. The third parameter is the cost of failing to respond to a target digit. In cell 3 the subject would lose 200 points, whereas in cell 2 only 10 points would be lost. We could also have manipulated the difficulty of response by making some cells four-choice reaction time cells and others only one-choice. However, the experiment that I will be telling you about had a four-choice reaction time task in each cell.

As you can see, this generic laboratory task which we call a visual attention allocation task contains the dimensions that we identified in the application set. We hope that it will provide a laboratory vehicle for investigating a number of theoretically important principles of performance. We believe that the task could be used to look at transfer of training



PARAMETER SET 1

1 .2 3 sec 0 pts. 0	2 .3 3 sec 10 pts. 1	3 4 2 sec 200 pts. 40
4 .2 2 sec 100 pts. 10	5 .4 4 sec 0 pts. 0	6 .1 3 sec 100 pts. 3.3

Figure 5. A parameter set for the Visual Attention Allocation Task.

variables, goal setting and knowledge of results, decision making strategies, performance on a single underload/overload continuum, and the affects of stress on performance.

To illustrate the use of this task, I will now describe an experiment in which we investigated a particular transfer of training principle. Of interest to us in this experiment was how the specificity of information provided during training affects performance during transfer to an unspecified parameter set. It might seem appropriate from a training point of view to provide the operator you are training with all of the relevant information required to perform the task maximally during training. However, such a strategy could fail to develop the operator's ability to discover the underlying parameter values when he or she is transferred to a situation not included in the training program. Imagine, for example, an aircraft pilot who has been trained to respond to a number of phases of flight. Suppose the aircraft malfunctions, the weather conditions change, or some other situation occurs for which the pilot has not been trained. The pilot must now formulate a new attention/allocation strategy to meet these conditions and must do this without the aid of additional training. Does the way the pilot was trained affect how quickly such an adjustment can be made?

In our experiment we contrast what we call our informed operator with a "learning-by-doing" operator. During training the informed operator is provided with all of the information required to formulate an appropriate allocation strategy. The learning-by-doing operator, on the hand, is simply given feedback about overall system performance and has to discover appropriate allocation strategies by using the system feedback to infer parameter values. The question of interest is whether operators trained using these two techniques differ in how quickly they can adopt an appropriate strategy when transferred to an unknown parameter set.

I have already described the visual attention allocation task to you. Two groups containing eight subjects each participated in the four phases of the experiment. For both groups, Phase I, which lasted for one hour on each two days, was used to familiarize the subjects with the mechanics of the task. Subjects were told that targets could occur in any of the six cells and when a target occurred, they should indicate the target location with the right hand and the target number with the left hand. During this phase neither group was given information about the underlying parameter set nor system feedback about the number of points earned on each trial. In each day of Phase I subjects were given four blocks containing five 2-minute trials. The only feedback given was a beep from the computer when an incorrect response was made.

Phase II started the actual training trials. At the beginning of Phase II the <u>informed</u> group was told the meaning of the three parameter values for each cell. Subjects in the informed group were also told the actual parameter values for each cell and in addition were given information about the expected loss in points per second. As you can see in Figure 5, this metric should have allowed subjects to determine the general importance of each cell, although it does not provide all the information that may be required to develop a optimal allocation strategy. Phase II lasted three days. On each day the subjects received six blocks of trials containing five 2-minute trials.

At the beginning of Phase II the learning-by-doing group was told about the parameter set dimensions that could be manipulated in each cell but were not told what the specific parameter values were. However, this group did receive the running point total at the bottom of the screen. Thus, this group received information about overall system performance but had no specific information about parameter set values.

At Phase III both groups were transferred to a new parameter set. Again the informed group was fully informed about parameter set values and expected points lost per second while the learning-by-doing group received only point total information.

Following the three days of Phase III training, both groups were transferred to Phase IV. During Phase IV only running point total information was given to the groups. The previously informed group no longer received information about parameter set values. The performance of the two groups during the early trials of Phase IV is of critical interest in this experiment. The question is whether the group trained using learning by doing would be superior to the informed group early in transfer.

While we have collected results concerning errors and latency of response for each of the cells during each trial, what I will present here is simply the average total points earned on each trial plotted for each block of trials.

During pretraining both groups improved from an average of -2350 points/trial to an average of -700 points/trial. By the end of pretraining the groups were performing comparably. Figure 6 shows the results of the training and transfer trials. During Phase II, as one might expect, the informed group was clearly superior to the learning-by-doing group and both groups showed considerable improvement in performance during this phase. At the start of Phase III the informed group showed no drop in performance and continued to improve throughout the phase. The learning-by-doing group, on the other hand, showed a small drop in performance in the beginning of Phase III, but then performance improved until at the end of this phase there was no difference in performance between the two groups.

At the beginning of Phase IV when transferred to an uninformed condition, the previously informed group showed a marked degradation in performance, dropping the same level of performance shown at the end of Phase I. The learning-by-doing group, on the other hand, showed a small drop in performance and quickly recovered. Rather surprisingly the informed group failed to catch up with the learning-by-doing group until the end of the third day.

The implication of these results seems clear. While the informed group out performed the learning-by-doing group early in training, after considerable training the groups performed equivalently. However, the learning-by-doing group apparently had learned a skill during training not learned by the informed group. This skill allowed the learning by doing group to more quickly determine and adopt an appropriate allocation strategy when put into a new situation having only system feedback. From a purely training point of view it might seem appropriate to provide as much information as possible to operators, but this strategy may be inappropriate should the operators be

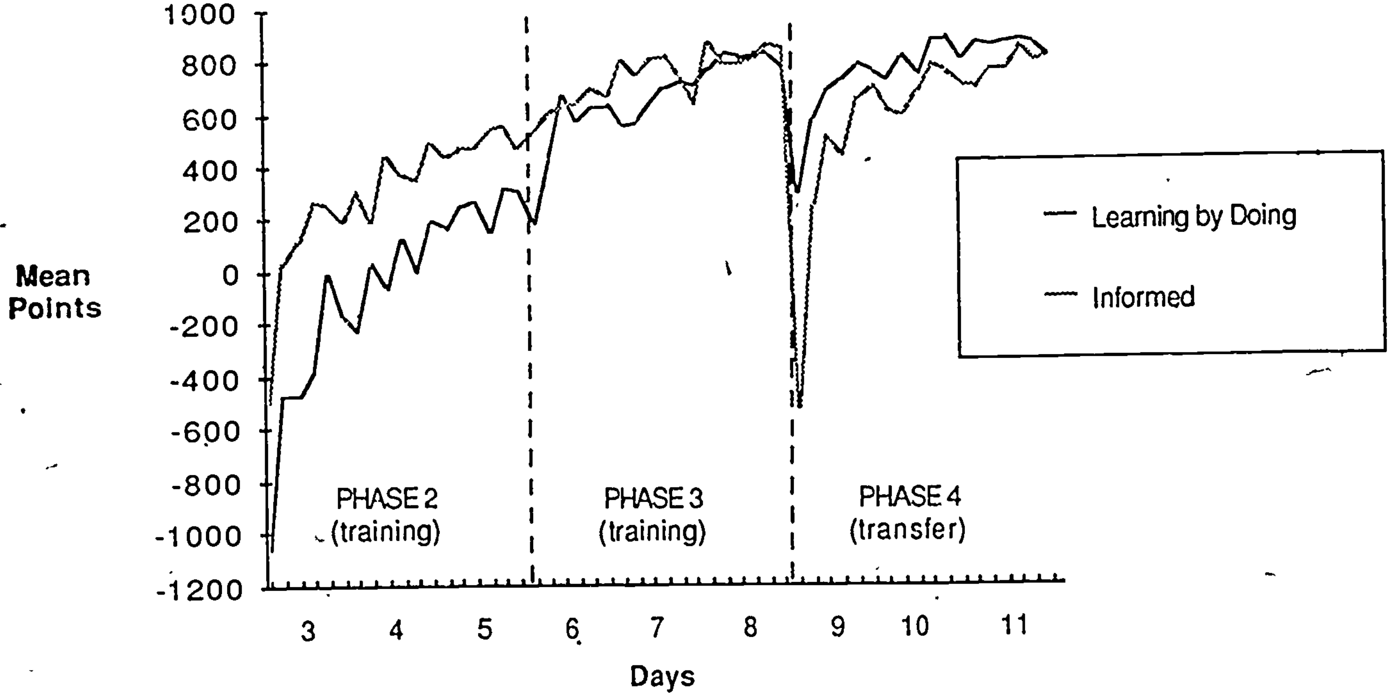


Figure 6. Mean points per trial earned by learning-by-doing subjects and informed subjects during training and transfer.

required to transfer to situations in which they have to adopt new strategies, such as during an emergency.

At this point we are to the theoretical conclusion stage of the suggested research strategy. Actually the conclusion that, "learning-by-doing training is superior when operators are transferred to new unknown system conditions" is not theoretically very rich. The conclusion needs to be elaborated on with explanatory features such as, "learning by doing is superior because operators learn how to derive the underlying system dynamics from system feedback." This elaboration can be tested using converging operations. For example, we could determine the operator's ability to specify the parameters of system dynamics. We also could test whether operators trained using learning by doing would be superior at using system feedback to indicate a change in system dynamics.

The final step in the suggested research strategy is to validate the theoretical conclusion in one or more application settings. In my laboratory we are currently attempting to validate the learning-by-doing conclusion. While I have no data to show you at this time, let me tell you a bit about our validation experiment.

The application setting that we have chosen is the aircraft pilot. Because we do not have access to an actual aircraft and because there are certain ethical considerations in allowing subjects to learn by doing when they crash a real aircraft, we have chosen to use a computerized flight simulator. We have implemented the Microsoft Flight Simulator on an IBM PC. The simulator is a sophisticated program which simulates the cockpit of a Cessna 182. The upper half of a color video display terminal shows the view out of a cockpit window with features on the ground and in the sky. The lower half of the screen shows the instrumentation. We are using two joy sticks for control. One of these joy sticks controls the ailerons and rudder using a side-to-side movement and the elevators using a forward-to-back movement. The other joy stick uses a forward-to-back movement to control the throttle setting. The dynamic response of the program has been rated as extremely good. We are using the simplest version of the program with coordinated rudder and ailerons and clear daylight flights with no clouds and no wind turbulence.

Basically we have designed three phases of flight which our subjects must learn. In one phase the aircraft is sitting on the runway and the subject must take off and attain straight and level flight at 1,000 ft. In a second phase the subject starts out in straight and level flight and must bank the aircraft in order to change the heading by 60 degrees. At the end of the maneuver the aircraft should be in straight and level flight with no change in in altitude or air speed. A third phase requires subjects to start in straight and level flight then descend 500 feet and change heading by 30 degrees. Again at the end of the maneuver the aircraft should be flying straight and level at the original airspeed.

The subject is given one minute to accomplish each of these phases of flight. At the end of each 1-min. trial the computer is "frozen" and the five parameters of altitude, heading, pitch, bank, and airspeed are determined and fed into a weighted formula which indicates system performance. The formula

produces numbers from 0 to 1,000 depending upon how close the subject was to achieving correct terminal performance. The parameters are weighted differently depending upon the particular phase of flight. For example, when the subject's task is to take off and attain a particular altitude, altitude is weighted heavily. When the subject's task is to bank the aircraft to achieve a particular heading, heading is weighted heavily. After each trial the subject is given the system feedback score for that trial.

We are currently training subjects in each of the three phases of flight using either learning-by-doing or informed training conditions. Informed subjects are told how to perform a particular maneuver. In particular, the instructions emphasize which instruments to pay attention to during that phase of flight in order to achieve good performance. The learning-by-doing subjects, on the other hand, are simply told what the instruments measure and how to read them. They are not told how to accomplish a particular phase of flight nor which instruments to pay attention to. They must derive this information from system performance and feedback. Each subject flys the computer for one hour a day, with 3 days devoted to each phase of flight. The first two phases are training phases and the third phase is transfer. During transfer all subjects are in a learning-by-doing mode. Obviously, we hope and expect that our basic finding from the generic laboratory task will be replicated using the flight simulator. If the learning-by-doing principle holds up and the results are similar to those generated from the generic laboratory task, then we will not only have replicated the theoretical conclusion, but we will have also gained confidence that we can generalize the conclusion to application settings, and we will have increased our confidence in the generic laboratory task as an appropriate testbed for our application set.

I hope that this case study has illustrated the usefulness of the suggested research strategy. I believe that a research strategy similar to the one suggested here is appropriate for those of us who wish to make a contribution to applied as well as basic areas of research. While this type of research requires more time and effort than using a simple laboratory paradigm, I believe that the ultimate contribution more than compensates for this added effort.